UNITED STATES PATENT APPLICATION

for

METHOD FOR SUPPORTING HIGH PRIORITY CALLS ON A CONGESTED WAN LINK

Inventor:

Prasad Miriyala

Prepared by:

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN
12400 WILSHIRE BOULEVARD
SEVENTH FLOOR
LOS ANGELES, CALIFORNIA 90025
(408) 720-8598

Attorney's Docket No. 81862.P164

"Express Mail" mailing label number EL371010874US
Date of Deposit October 15, 1999
I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231.

Patricia A. Balero
(typed or printed name of person mailing paper or fee)

(Signature of person mailing paper or fee)

METHOD FOR SUPPORTING HIGH PRIORITY CALLS ON A CONGESTED WAN LINK

Field of the Invention

This invention relates to resource allocation within a computer or other network. More particularly, the present invention relates to the allocation of bandwidth in a Digital Communications Network.

Background

Figure 1 generally depicts a digital network **100** adapted to transport voice, data and other information. This exemplary network will be used to illustrate some of the drawbacks associated with present networking communication methodologies and equipment. The network includes a plurality of nodes **102, 104** (hereinafter "network nodes"), each of which may be coupled to various customer premises equipment (CPE) (not shown) and/or other network nodes (not shown). Network nodes **102** and **104** are communicatively coupled to one another through a communication link **106**, which may be a wired or wireless communication link. In this network **100**, nodes **102** and **104** communicate across communication link 106 using the well-known Asynchronous Transfer Mode (ATM) protocol. Thus, at each end of communication link **106**, ATM interfaces **108** are employed. Communication between nodes 102 and 104 with various CPE and/or other nodes is supported across communication links **108** and **110**, respectively, each of which is associated with an appropriate interface **112, 114** at its respective node.

When a communication (e.g., a voice communication) is to be transported within network **100**, the signal first travels from the associated CPE equipment (e.g., a private branch exchange or PBX) to node **102** over communication link (e.g., digital T1 carrier) **108**. Before being transmitted over such a T1 carrier, the signal is sampled and converted to a digital signal. A common sampling rate used with voice communications is 8000 samples per second, with each digital sample represented by 8 bits of data. Thus, the data rate of the new digital signal is: 8000 samples/sec x 8 bits = 64,000 bits/sec. This technique is known as Pulse Code Modulation (hereinafter "PCM") and is used extensively throughout the backbone of modern telephone systems.

Although no international standard has been adopted, the T1 carrier is one method of PCM used throughout North America and Japan. The T1 carrier is comprised of 24 channels of digital data multiplexed together. Digitally sampled data from each of the 24 channels are packaged into successive frames of 8 bits/channel x 24 channels + an additional framing bit = 193 bits. Outside of North America and Japan a similar standard, known as E1, is commonly implemented. E1 operates in a manner similar to T1 except that it uses 32, 8-bit data samples (i.e., 32 channels) instead of 24.

After the signal has been sampled, converted to a digital signal and transmitted over a T1 carrier, as described above, it is then transferred to an outgoing communication link **106**. Because the transport protocol across communication link **106** is different than that used on communication link 108, the digital data samples are packaged according to the protocol used across communication link **106** (e.g., ATM) before being transmitted to node **104**. Additionally, although PCM by itself provides for a data transfer rate of 64,000 bits/second, it is often desirable to further compress the digital PCM data in order

to save bandwidth within the network. This can be accomplished using Digital Signal Processing (hereinafter "DSP") resources associated with network node **102**. For example, if the 64,000 bits/second PCM signal is compressed by a DSP resource at a compression ratio of 16:1, the resulting digital signal will be transmitted at 4,000 bits/second. This represents a significant reduction in required bandwidth across network **100** to transmit the same underlying signal. Such compression techniques are particularly useful in networks that are heavily loaded with network traffic. Examples of compression algorithms known in the art include the International Telegraph Union (hereinafter "ITU') standards G.711, G.726, G.729-A, G.729, and G.728. Such compression resources may be associated with the ATM interfaces **108** and may operate under the control of a node controller **116** in each of the nodes.

Sub a1

However, there is a tradeoff between bandwidth savings over network **100** and the implementation of costly DSP resources at the nodes. In general, the higher the compression ratio required by the compression algorithm, the more DSP resources are used up processing the compression request over a given period of time. Thus, while a single DSP resource may process up to say 16 channels of data if no compression is used (i.e., in baseline PCM mode), it may be limited to 5 channels if data is compressed at 2:1, and only 2 channels if the PCM signal is compressed at 8:1. One factor behind this limitation is the limited period of time in which the DSP resource must operate compress the data within a T1 frame before it must move on to the next frame of data. Thus, the chosen compression ratio will have a significant impact on DSP resource usage.

Sub a2

Following compression (if used), the data samples are delivered through ~~network~~ 100 ~~to node~~ **104**, where the data may be decompressed and passed on to other CPE or

another node. The system is bi-directional to ensure 2-way communication ~~between the nodes.~~

Sub A3

One problem with the communication scheme adopted in network 100 occurs when communication link 106 becomes congested, that is, when there is no available bandwidth to support new incoming calls from CPE coupled to node **102**. Consider, for example, a situation where multiple calls being transported between nodes **102** and **104** are using all or almost all of the available bandwidth on communication link **106**. If a high priority call (e.g., a 911 or other emergency call) is now received at network node 102, either of two scenarios is possible. First, the high priority call may be rejected (dropped) in the face of no available bandwidth. Second, rather than dropping the high priority call (clearly a least acceptable solution); the nodes may be configured to drop lower priority calls in order to free up bandwidth to accommodate the high priority call. Although this solution may allow the high priority call to proceed, it is less than satisfactory in as much as several existing calls may be dropped to support the one new call. What is needed, therefore, is a more robust mechanism for handling such situations.

Summary of the Invention

Sub 6 a4

~~. In one embodiment, a network node is configured to negotiate for connections for high priority calls (e.g., voice calls) received at the node in the face of otherwise congested outbound communication links. The negotiation is conducted in a fashion that will preserve connections for existing calls associated with the node. For example, the negotiation may be conducted so as to cause one or more of the existing calls to consume less bandwidth over the outbound communication links than was consumed at a time prior to reception of the high priority calls. Such negotiations may be initiated depending on the availability of codec resources and/or compression schemes at the node.~~

Another embodiment provides a method of managing a communication link between nodes of a communication network so as to ensure connection availability for one or more high priority calls over the communication link through dynamic renegotiations of call parameters for existing calls (e.g., voice calls) transported over the communication link. The communication link may preferably support communications according to the Asynchronous Transfer Mode and the dynamic renegotiations may be negotiations of compression schemes for the calls. Such dynamic renegotiations may be supported according to codec availability (e.g., as determined according to profile information) at the nodes and may be accomplished through the exchange of OAM cells between the nodes. In such a scheme, the high priority calls may be determined as such according to database information regarding called numbers.

Still another embodiment provides a network having a number of nodes connected through one or more communication links and a resource manager configured to allocate bandwidth over the communication links to high priority calls received at one or more of

the nodes without dropping existing calls within the network. The resource manager allocates bandwidth through dynamic renegotiations of existing bandwidth utilization within the network and may be a distributed resource among the nodes of the network. Preferably, the nodes each support multiple codec resources, which compress voice information transmitted over the communication link. The dynamic renegotiations are supported through the exchange of OAM cells between the nodes.

Brief Description of the Drawings

The present invention is illustrated by way of example and not limitation in the figures of the accompanying drawings, in which:

FIG. **1** shows generally a data network that can be configured in accordance with the present invention.

FIG. **2** shows an example of a process for renegotiating compression or other call parameters on the fly in accordance with an embodiment of the present invention.

FIG. **3** shows an example of a database for mapping incoming calls to outbound channels at a node of a communication network in accordance with an embodiment of the present invention.

Detailed Description

Described herein is a method for accommodating high priority calls on a congested communication link of a wide area or other communication network. In the following discussion, examples of specific embodiments of the present invention are set forth in order to provide the reader with a through understanding of the present invention. However, many of the details described below can be accomplished using equivalent methods or apparatus to those presented herein. Accordingly, the broader spirit and scope of the present invention is set forth in the claims that follow this detailed description and it is that broader spirit and scope (and not merely the specific examples set forth below) that should be recognized as defining the boundaries the present invention.

The basic mechanism adopted in the present invention is easily understood with reference to **Figure 2**. Process **200** describes a bandwidth renegotiation scheme for accommodating high priority calls over otherwise congested communication links in a wide area or other communications network. When a high priority call is received at a node of the network, a check is made to determine whether the call can be accommodated within the existing bandwidth utilization in the network.

For example, in the network **100** illustrated in **Figure 1**, if a high priority call is received at node **102**, controller **116** may be configured to determine whether the call can be accommodated on communication link **106** according to the existing bandwidth utilization thereof. If the call can be so accommodated, it is connected in the customary fashion. If the call cannot be so accommodated, however, the present bandwidth renegotiation scheme is employed to free up bandwidth on communication link **106** to the

point where the high priority call can be accommodated. Rather than simply dropping calls, by renegotiating the bandwidth utilization among existing calls, these calls as well as the new high priority call can be accommodated within network **100**.

sub Q5

To more fully appreciate the processes involved in the present scheme, it is helpful to understand how calls are handled in network **100** in accordance with the present invention. When a call is received from a CPE at node 102, it is mapped to an associated network address. For example, associated with controller 116 may be a database configured to provide appropriate mappings between dialed telephone numbers and network (e.g., ATM or Internet protocol (IP) addresses. An example of such a database **300** is shown in **Figure 3**. In this example, the telephone number (123) 456-7891 is mapped to the network address 123.221.456.78.

For the case of ATM network 100, this network address is associated with a virtual circuit (e.g., a permanent virtual circuit or PVC) between network nodes **102** and **104**, supported on communication link **106**. Thus, when the incoming call is parsed at node **102**, the dialed number may be extracted to determine the node for which the call is destined, according to the network address. In some cases, the dialed number and/or associated network address may be flagged as a high priority call. For example, the dialed number 911 may be flagged as a call of the highest possible priority (e.g., where more that two priority levels are available). In this way, incoming calls can be recognized as high priority or not.

As mentioned above, for voice calls a variety of compression schemes are available, depending upon the codec (coder-decoder) resources available at the end points of the communication link over which the call is transported. Each call (e.g., each

inbound TI channel at node 102) may negotiate for a particular compression scheme to be employed at the time a connection is established, or a default compression scheme may be used where no negotiation takes place. In accordance with the present invention, each voice port supported by the ATM interfaces **108** has a profile defining the available codec resources for that port. Incoming calls are mapped to these voice ports for communications across communication link **106** and in one embodiment, ATM interfaces **108** may each have up to 24 voice ports.

The profiles of the voice ports may be established by a network manager at the time PVCs are set up within network 100. Alternatively, or in addition, profiles may be exchanged between ports as part of a call set up process in the case of switched virtual circuits (SVCs). On mechanism for the exchange of such profiles is the ATM Adaptation Layer type 2 (AAL2) protocol. Recently, the ATM Forum has promulgated standards document af-vota-0113.000, entitled "ATM Trunk Networking Using AAL2 for Narrowband Services" (Feb. 1999). In that document, which is incorporated herein by reference as is set forth in its entirety, a scheme for selecting and managing encoding algorithms at nodes of an ATM network according to prearranged agreements is described. This scheme calls for the exchange of profile information in a manner suitable for use in accordance with the present invention.

By exchanging profiles, each node at an end of a communication link is aware of the codec (e.g., DSP) resources available at each corresponding voice port. Thus, the nodes can reach an agreement on which codec (i.e., which compression scheme) to use for a particular call. As noted above, compression schemes that provide high compression ratios tend to utilize less bandwidth than those that use low compression

ratios do. So, assuming that sufficient codec resources are available, when an incoming high priority call is recognized, if there is insufficient bandwidth available on communication link **106** to accommodate the new call, node **102** can instruct node 104 to adopt new compression schemes on one or more voice channels (i.e., calls) so as to free up bandwidth on communication link **106** to support the new call. This process can be repeated, as required, until no available codec resources remain.

The renegotiation process described above may be implemented using OAM (operations, administration and maintenance) cells that are exchanged between nodes 102 and 104. OAM cells are often exchanged between nodes of an ATM network and are used to convey a variety of information. In accordance with the present invention, the OAM cells are configured with payloads (e.g., cell type, function type and/or function specific fields) that contain instructions for moving to different compression schemes, and acknowledgements thereto. The format and use of OAM cells are well known in the art and need not be further described herein. What is unique is the use of such cells (which may be transmitted from a high priority queue within a node so as to ensure rapid call handling) to negotiate compression schemes for a channel on the fly.

Sub a6

Thus, a scheme for has been described. Although discussed with respect to specific embodiments, however, the broader applicability of the present invention should not be limited thereby. For example, although discussed with respect to the negotiation of compression schemes on the fly, other call parameters or connection parameters could be so negotiated on the fly. Thus, this broader applicability of the present invention is recited in the claims that follow.